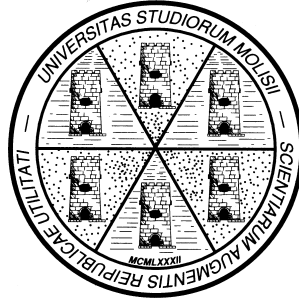


Università degli Studi del Molise  
Facoltà di Economia  
Dipartimento di Scienze Economiche, Gestionali e Sociali  
Via De Sanctis, I-86100 Campobasso (Italy)



ECONOMICS & STATISTICS DISCUSSION PAPER  
No. 14/03

**A possibilistic approach to latent structure analysis  
for symmetric fuzzy data**

by

Pierpaolo D'Urso  
*University of Molise, Dept. SEGeS*

and

Paolo Giordani  
*“La Sapienza” University, Rome. Dept of Statistics, Probability, and Applied Statistics*

The Economics & Statistics Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. The views expressed in the papers are solely the responsibility of the authors.

# A possibilistic approach to latent structure analysis for symmetric fuzzy data

Pierpaolo D'Urso

*Dipartimento di Scienze Economiche, Gestionali e Sociali,  
Università degli Studi del Molise, Via De Sanctis, 86100 Campobasso, Italy.*

E-mail: [URSO@unimol.it](mailto:URSO@unimol.it), [pierpaolo.durso@uniroma1.it](mailto:pierpaolo.durso@uniroma1.it).

Paolo Giordani

*Dipartimento di Statistica, Probabilità e Statistiche Applicate,  
Università degli Studi di Roma "La Sapienza", P.le A. Moro, 5 - 00185 Roma, Italy.*

E-mail: [paolo.giordani@uniroma1.it](mailto:paolo.giordani@uniroma1.it).

**Abstract:** In many situations the available amount of data is huge and can be intractable. When the data set is single valued, latent structure models are recognized techniques, which provide a useful compression of the information. This is done by considering a regression model between observed and unobserved (latent) fuzzy variables. In this paper, an extension of latent structure analysis to deal with fuzzy data is proposed. Our extension follows the possibilistic approach, widely used both in the cluster and regression frameworks. In this case, the possibilistic approach involves the formulation of a latent structure analysis for fuzzy data by optimization. Specifically, a non-linear programming problem in which the fuzziness of the model is minimized is introduced. In order to show how our model works, the results of two applications are given.

**Keywords:** Latent structure analysis, symmetric fuzzy data set, possibilistic approach.

## 1. Introduction

There are many real situations in which the observed data are directly fuzzy or a process of fuzzification is required. In fact, several fields are the domain of these type of datasets. For instance, fuzzy data are widely analyzed in: ballistics (Celminš, 1991), event tree analysis (Huang et al., 2001), food chemistry (Kallithraka et al., 2001), group consensus opinion and multicriteria decision making (Herrera et al., 1997; Lee, 2002; Raj and Kumar, 1999), human errors rate assessment (Richei et al., 2001), machine learning (Chung and Chan, 2003), management (managerial talent assessment) (Chang et al., 2000), maritime safety (based on decision making) (Sii et al., 2001), material selection analysis (Chen, 1997), medical diagnosis (Di Lascio et al., 2002; Kuncheva and Steimann, 1999; Steimann and Adlassnig, 1994), web advertising (Coppi and D'Urso, 2003), military application (Cheng and Lin, 2002; Cheng et al., 1999), nuclear energy (Moon and Kang, 1999), intelligent manufacturing (Shen et al., 2001), petrophysics (Finol et al., 2001), process control (Laviolette et al., 1995), risk analysis (Lee, 1996), car performances (D'Urso, 2003), restaurant performances (D'Urso, 2003) software reliability (D'Urso and Gastaldi, 2002), technical efficiency (Hougaard, 1999), thermal sensation analysis (Hamdi et al., 1999), VDT legibility (Chang et al., 1996), cement composition (Xu and Li, 2001). For example, in medical diagnosis, Di Lascio et al. (2002), for classifying two samples of patients, hospitalized in different times with diabetic neuropathy, according to the severity of the symptoms, proposed suitable linguistic labels denoting a different severity grade of symptom. They represent each label by a triangular fuzzy number. The values are assigned partly by consultancy of an expert in the fields, and partly by activating the phase of initial training.

In all the above cases, the data at hand are massive and, thus, not easily tractable in their original dimensions. In order to compress the available data and to extract relevant features, latent structure models are well known tools, which synthesize huge amounts of single valued data losing relevant information as little as possible. Suppose to deal with  $I$  observation units characterized by  $J$  (single valued) variables. Each observation unit can be represented as a point in the reference space  $\mathfrak{R}^J$ . In this respect,  $P < J$  unobserved variables, which are linear combinations of the observed ones, are extracted. Usually the number of unobserved variables,  $P$ , is chosen equal to 2. In this way, each observation unit is now represented as a point in the reference space  $\mathfrak{R}^P$ . Therefore,

when  $P = 2$ , we get a bi-dimensional configuration of the  $I$  objects, which can be easily managed.

The extracted latent variables maintain relevant features, which underlie the data. This is based upon the empirical assumption that the observed variables are strongly related to theoretical variables, which are not observable. In this respect, a latent structure model can be seen as a particular regression analysis in which the explanatory variables are unobservable, and, thus, unknown. For example, in marketing research, we may consider the opinion and the behaviour of respondents concerning the economic life of a country. The available data about respondent's opinion and behaviour are related to the political point of view of the respondents, which is unobservable. Similarly, we may wish to study theoretical variables such as quality of life, wellbeing, intelligence, knowledge, which are not observable but are related to several observed variables. The data set in Table 1 (Dencœux and Masson, 2003) may help to clarify this assumption.

**Table 1:** Student data

Student	Mathematics 1	Mathematics 2	Physics 1	Physics 2
Tom	15	Fairly good	Unknown	[14,16]
David	9	Good	Fairly good	10
Bob	6	[10,11]	[13,20]	Good
Jane	Fairly good	Very good	19	[10,12]
Joe	Very bad	Fairly bad	[10,14]	14
Jack	1	[4,6]	9	[6,9]

Note: the intervals are expressed as [lower bound, upper bound].

Table 1 refers to marks obtained by six students in mathematics and physics. Even if the Student data set is relatively small, it is difficult to find a rating among the students. In fact, by inspecting Figure 1, it is not easy to evaluate the knowledge of the involved students. In this case, it is plausible to assume that the marks in mathematics and physics are related to the knowledge of the students in mathematics and physics, respectively. Thus, two latent variables may underlie the four observed variables. A rating among the students can then be easily obtained by considering the bi-dimensional representation in the reference space  $\mathfrak{R}^2$  where the two axes can be interpreted as the

knowledge in each matter. Suppose that the number of students as well as that of scholastic matters increases. For instance, suppose that the matters are literature, history, philosophy, mathematics, physics, chemistry. In this case, by extracting two latent variables, it may occur that one latent variable can be interpreted as the arts learning (first three matters) and the latter one as the scientific learning (latter three matters).

Unfortunately, standard techniques deal with crisp (non-fuzzy) data. Therefore, they are not able to handle data sets as that given in Table 1. In this paper, we propose an extension of classical latent structure models when the observed data are fuzzy following a possibilistic approach. This helps us to find the underlying latent structure of observed fuzzy data by considering all the available information (centers, spreads and membership functions) that may be arbitrarily overlooked when dealing with the centers only. The use of all the available information offers a deeper insight into the phenomenon at hand and avoids that the performed analysis gives possible misleading results. The possibilistic theory plays an important role in many real world situations for reflecting the ambiguity and vagueness of human understanding, the incompleteness of information, cognitive ignorance and diversity of evaluation. In all these cases, the available information is vague and therefore cannot be revealed exactly by crisp data. See, for more details, Dubois and Prade (1980); Nahmias (1978).

The possibilistic approach can be considered in different statistical data analyses. In particular, possibilistic theory is diffusely utilized in cluster analysis (see, for instance, Krishnapuram and Keller, 1993; Barni et al., 1996; Krishnapuram and Keller, 1996; Mènard et al., 2003) as well as in regression analysis (see, for instance, Tanaka et al., 1980, 1982; Chang et al., 1996; Tanaka and Lee, 1998; Chen, 2001; Lee and Chen, 2001). In particular, in the regression framework, “possibilistic regression is formulated from the possibility viewpoint to obtain the smallest interval system containing all the selected data” (Tanaka et al., 1982).

In this work, we utilize the possibilistic approach in order to obtain the smallest interval system containing the data while extracting their latent components. A few extensions of latent structure models for fuzzy data are available in the literature. Yabuuchi et al. (1997) and Watada and Yabuuchi (1997) propose a method to perform Principal Component Analysis on fuzzy data, in which fuzzy eigenvalues and crisp eigenvectors are obtained by solving  $P$  linear programming problems but inclusion constraints are

not considered. Giordani and Kiers (2004) suitably generalize PCA to deal with fuzzy data by following a least-squares approach. Denceux and Masson (2003) propose a generalization of PCA for fuzzy data using autoassociative neural networks. Finally, Factor Analysis was generalized to fuzzy data by Nakamori and Watada (1997).

The paper is structured as follows. In Section 2, we synthetically recall classical latent structure models; successively, in Section 3, we define the notion of symmetric fuzzy data and propose our generalization of latent structure models for symmetric fuzzy data sets. Then, in Section 4, two illustrative applications to are shown.

## 2. Latent structure models

Let us suppose to deal with  $I$  observation units characterized by  $J$  crisp variables. In order to compress such a data set losing relevant information as little as possible Principal Component Analysis is probably the most used statistical exploratory tool. To do so,  $P < J$  unobserved variables, called components, which are linear combinations of the observed ones, are extracted. The latent structure model can then be described as

$$\mathbf{X} = \mathbf{A}\mathbf{F}' + \mathbf{E} \quad (1)$$

where  $\mathbf{A}$  is the component score matrix of order  $(I \times P)$ ,  $\mathbf{F}$  is the component loading matrix of order  $(J \times P)$  and  $\mathbf{E}$  is the residual matrix. Each column of  $\mathbf{A}$  and  $\mathbf{F}$  refers to one latent variable. The matrix  $\mathbf{F}$  contains the coefficients of the linear combinations of the observed variables, which define the unobservable variables. The matrix  $\mathbf{A}$  contains the regression parameters that relate the observed variables to the latent ones. Since  $\mathbf{F}$  is columnwise orthonormal, it follows that the latent variables are uncorrelated.

The optimal solution is obtained by minimizing, in the least-squares sense, the sum of squared residuals:

$$\|\mathbf{X} - \mathbf{A}\mathbf{F}'\|^2 \quad (2)$$

where  $\|\cdot\|^2$  denotes the Euclidean norm, that is the sum of squares of the elements of the matrix involved<sup>1</sup>. In order to choose the number of extracted components  $P$ , taking into account that the solutions are nested, one usually considers the point such that extracting one additional component leads to a negligible decrease of the sum of squared residuals.

The solution is obtained by means of the Singular Value Decomposition (SVD) of  $\mathbf{X}$ . It provides the best  $P$ -rank decomposition of  $\mathbf{X}$  as  $\mathbf{P}_p \mathbf{D}_p \mathbf{R}_p'$  where  $\mathbf{P}_p$ , of order  $(I \times P)$ , and  $\mathbf{R}_p$ , of order  $(J \times P)$ , are matrices containing the first  $P$  unit length (left and right) singular vectors of  $\mathbf{X}$  and  $\mathbf{D}_p$  is the diagonal matrix displaying on the diagonal the first  $P$  singular values of  $\mathbf{X}$ . It follows that the solution is

$$\mathbf{A} = \mathbf{P}_p \mathbf{D}_p, \quad (3)$$

$$\mathbf{F} = \mathbf{R}_p. \quad (4)$$

The solution is usually interpreted by considering the loading matrix  $\mathbf{F}$ . In fact, high loadings indicate strong relations between the original variables and the latent variable at hand. To improve the interpretation, it can be useful to plot the entities of one mode in the low-dimensional space spanned by the columns of the other mode. For example, the visualization of the observation units can be done by using the scores in  $\mathbf{A}$  as coordinates provided that the loading matrix  $\mathbf{F}$  is columnwise orthonormal in order to have an adequate plot. Notice that the estimated data matrix is the orthogonal projection (with coordinates in  $\mathbf{A}$ ) of the original data matrix onto the particular subspace spanned by the columns of  $\mathbf{F}$ .

### 3. Latent structure models for fuzzy data

#### 3.1. Fuzzy data

Preliminarily, we define the notion of symmetric fuzzy data. A generic fuzzy datum is formalized as  $\tilde{x}_{ij} = (m_{ij}, s_{ij})_L : i = 1, \dots, I; j = 1, \dots, J$ , where  $m_{ij}$  and  $s_{ij}$  are,

---

<sup>1</sup> Notice that, given a matrix  $\mathbf{Y}$ ,  $\|\mathbf{Y}\|^2 = \text{tr}(\mathbf{Y}'\mathbf{Y}) = \text{tr}(\mathbf{Y}\mathbf{Y}')$ .

respectively, the *center* and the (left and right) *spread* of the  $j$ -th fuzzy variable observed on the  $i$ -th object with the following membership function:

$$\mu_{\tilde{X}_{ij}}(u_{ij}) = L\left(\frac{u_{ij} - m_{ij}}{s_{ij}}\right), \quad s_{ij} > 0, \quad (5)$$

where  $L(z)$  indicates a decreasing *shape function* from  $\mathfrak{R}^+$  to  $[0,1]$  with  $L(0)=1$ ;  $L(z)<1$  for all  $z>0$ ;  $L(z)>0$  for all  $z<1$ ;  $L(1)=0$  (or  $L(z)>0$  for all  $z$  and  $L(+\infty)=0$ ) (Dubois and Prade, 1980).

Very interesting classes of symmetric fuzzy data are characterized by the following families of membership functions (Tanaka and Guo, 1999):

$$\begin{aligned} L\left(\frac{u_{ij} - m_{ij}}{s_{ij}}\right) &= \max\left\{0, 1 - \left|\frac{u_{ij} - m_{ij}}{s_{ij}}\right|^q\right\}, \\ L\left(\frac{u_{ij} - m_{ij}}{s_{ij}}\right) &= \exp\left(-\left|\frac{u_{ij} - m_{ij}}{s_{ij}}\right|^q\right), \\ L\left(\frac{u_{ij} - m_{ij}}{s_{ij}}\right) &= \left(1 + \left|\frac{u_{ij} - m_{ij}}{s_{ij}}\right|^q\right)^{-1}, \end{aligned} \quad (6)$$

with  $s_{ij} > 0, q > 0$ . Particular cases of the first family of membership function, are the

*square root* membership function for  $q = \frac{1}{2}$ , the *symmetric triangular* membership

function for  $q = 1$ , the *parabolic* membership function for  $q = 2$ .

In several real life applications, a researcher may deal with a  $J$ -vector of fuzzy numbers. In order to define the membership function of a  $J$ -dimensional fuzzy number, at least two proposals can be adopted. Celminš (1987) assumes that the  $J$ -dimensional fuzzy points are hyperellipsoids in  $\mathfrak{R}^J$  whose membership function depends on a vector called *apex* and a positive definite matrix called *panderance matrix*. The concept of *non-interactivity* (Hisdal, 1978) among fuzzy numbers can also be followed. In this respect, the  $J$ -dimensional fuzzy numbers are hyperrectangles whose total number of vertices is  $2^J = K$ , in which the distribution of the points depends on the  $J$  membership

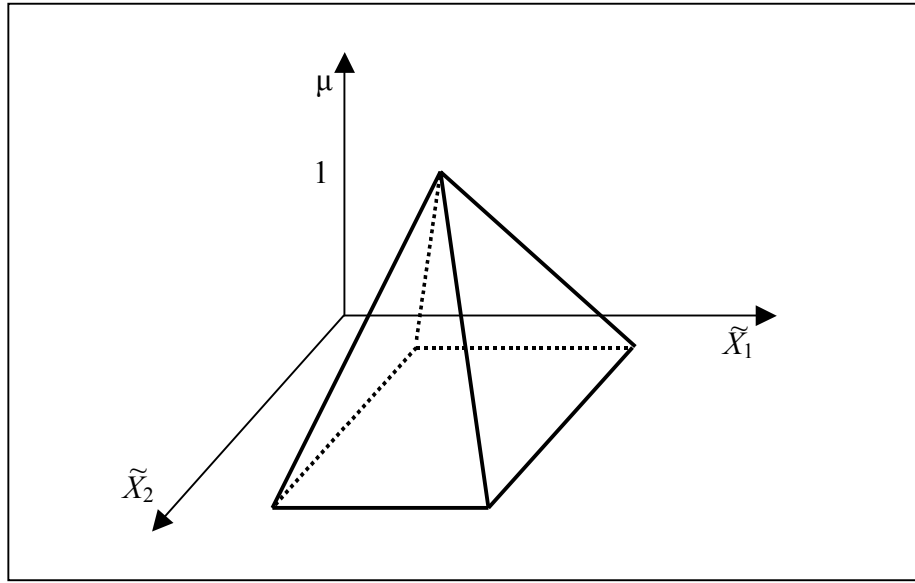


functions. Let  $\tilde{\mathbf{X}} = (\tilde{X}_1, \dots, \tilde{X}_J)$  be a  $J$ -dimensional non-interactive fuzzy vector (see Figure 1). The membership function of  $\tilde{\mathbf{X}}$  is

$$\mu_{\tilde{\mathbf{X}}}(\mathbf{u}) = \min_{j=1, \dots, J} \mu_{\tilde{X}_j}(u_j). \quad (7)$$

Notice that this proposal, which is considered in this paper, does not involve interrelations among the  $J$  fuzzy variables.

**Figure 1:** Non-interactive fuzzy numbers ( $J = 2$ ).



### 3.2. The model

In the regression framework, the optimal solution is obtained by solving a linear programming problem (LP) in which the fuzziness of the model is minimized and the inclusion relations between the estimated and observed fuzzy output variable are the constraints. Since each observation unit is represented by a score on a single fuzzy variable (the output variable), the inclusion constraints refer to the bound of fuzzy intervals in  $\Re^1$ . In the data reduction framework, since each observation unit is represented by a hyperrectangle in  $\Re^J$ , the inclusion constraints refer to all the  $K$  vertices that characterize each hyperrectangle.

Let  $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \dots, \tilde{x}_{iJ})$ ,  $i = 1, \dots, I$ , be  $J$ -dimensional symmetric fuzzy vectors stored into the data matrix  $\tilde{\mathbf{X}}$  of order  $(I \times J)$  and let  $\mathbf{M}$  and  $\mathbf{S}$  be, respectively, the matrices of the

centers and of the spreads. Following a possibilistic approach, the latent structure model for fuzzy data can be expressed as

$$\tilde{\mathbf{X}} = \tilde{\mathbf{A}}\mathbf{F}' \quad (8)$$

where  $\tilde{\mathbf{A}} = (\mathbf{A}_M, \mathbf{A}_S)_L$  is the fuzzy component score matrix of order  $(I \times P)$  whose generic element is the symmetric fuzzy number denoted as  $(a_{Mip}, a_{Sip})_L$ . Let  $\mathbf{A}_M$  and  $\mathbf{A}_S$  be, respectively, the matrices of the centers and of the spreads whose generic elements are  $a_{Mip}$  and  $a_{Sip}$ ,  $i = 1, \dots, I$ ;  $p = 1, \dots, P$ . Finally,  $\mathbf{F}$  is the crisp component loading matrix of order  $(J \times P)$ . Therefore, in our possibilistic extension, we assume that the component score matrix is fuzzy in order to take into account the fuzziness of the data involved. In this respect, the fuzzy possibilistic latent structure model can be seen as a particular fuzzy possibilistic regression such that the matrix  $\mathbf{A}$  provides the coefficients that relate the  $J$  observed fuzzy variables to the  $P$  unobservable crisp latent variables which can be defined by considering the columns of  $\mathbf{F}$ . Since the latent variables are uncorrelated (as we will see, we shall impose the columnwise orthonormality of  $\mathbf{F}$ ), the obtained possibility distributions of the coefficients are non-interactive.

Let us assume the following in order to formulate the latent structure model:

- i) The observed data  $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \dots, \tilde{x}_{iJ})$ ,  $i = 1, \dots, I$ , can be represented by model (8).
- ii) Given a threshold  $h$ , the observed data should be included in the  $h$ -level set of the estimated data. Given a fuzzy number  $\tilde{X}$ , the  $h$ -level set is  $[\tilde{X}]_h = \{x | \mu_{\tilde{X}}(x) \geq h\}$ <sup>2</sup>. As in possibilistic regression, the term  $h$  can be considered as a measure of goodness of fit (see, for details, Tanaka and Guo, 1999). If  $h=0$ , the estimated data totally enclose the observed ones.

---

<sup>2</sup> The strong  $h$ -level set can be also defined. Given a fuzzy number  $\tilde{X}$ , the  $h$ -level set is  $(\tilde{X})_h = \{x | \mu_{\tilde{X}}(x) > h\}$ .

iii) The index of the spread of the latent structure model is defined by

$$J = \sum_{i=1}^I \sum_{j=1}^J \mathbf{a}_{Si} |\mathbf{f}'_j|. \quad (9)$$

where the vector  $\mathbf{a}_{Si}$  is the  $i$ -th row of  $\mathbf{A}_S$ ,  $i=1, \dots, I$ , and  $|\mathbf{f}'_j|$  denotes the  $j$ -th row of  $\mathbf{F}$ , whose elements are in absolute value.

The well known fuzzy operations yield the generic estimated datum as follows,

$$\tilde{x}_{ij}^* = (\mathbf{a}_{Mi} \mathbf{f}'_j, \mathbf{a}_{Si} |\mathbf{f}'_j|)_L, \quad i=1, \dots, I, \quad j=1, \dots, J, \quad (10)$$

where the vector  $\mathbf{a}_{Mi}$  is the  $i$ -th row of  $\mathbf{A}_M$ ,  $i=1, \dots, I$ . From (10), it also follows that we get the estimated fuzzy data matrix of order  $(I \times J)$

$$\tilde{\mathbf{X}}^* = (\mathbf{M}, \mathbf{S})_L = (\mathbf{A}_M \mathbf{F}', \mathbf{A}_S |\mathbf{F}'|)_L. \quad (11)$$

The membership function of the generic estimated datum  $\tilde{x}_{ij}^*$  is

$$\mu_{\tilde{x}_{ij}^*}(x) = L((x - \mathbf{a}_{Mi} \mathbf{f}'_j) / \mathbf{a}_{Si} |\mathbf{f}'_j|), \quad i=1, \dots, I, \quad j=1, \dots, J, \quad (12)$$

for  $\mathbf{f}_j \neq \mathbf{0}$  where  $\mathbf{a}_{Mi}$  and  $\mathbf{a}_{Si}$  are, respectively, the centers and the spreads of the fuzzy component scores pertaining to the  $i$ -th observation unit. If  $\mathbf{f}_j = \mathbf{0}$  and  $x=0$ , we set  $\mu_{\tilde{x}_{ij}^*}(x)=1$  and, if  $\mathbf{f}_j = \mathbf{0}$  and  $x \neq 0$ ,  $\mu_{\tilde{x}_{ij}^*}(x)=0$ . From (12), we can obtain the  $h$ -level set of the generic estimated datum. With respect to the  $i$ -th observation unit and the  $j$ -th variable, the  $h$ -level set is

$$[\tilde{x}_{ij}^*]_h = [\mathbf{a}_{Mi} \mathbf{f}'_j - |L^{-1}(h) \mathbf{a}_{Si} |\mathbf{f}'_j|, \mathbf{a}_{Mi} \mathbf{f}'_j + |L^{-1}(h) \mathbf{a}_{Si} |\mathbf{f}'_j|], \quad i=1, \dots, I, \quad j=1, \dots, J. \quad (13)$$

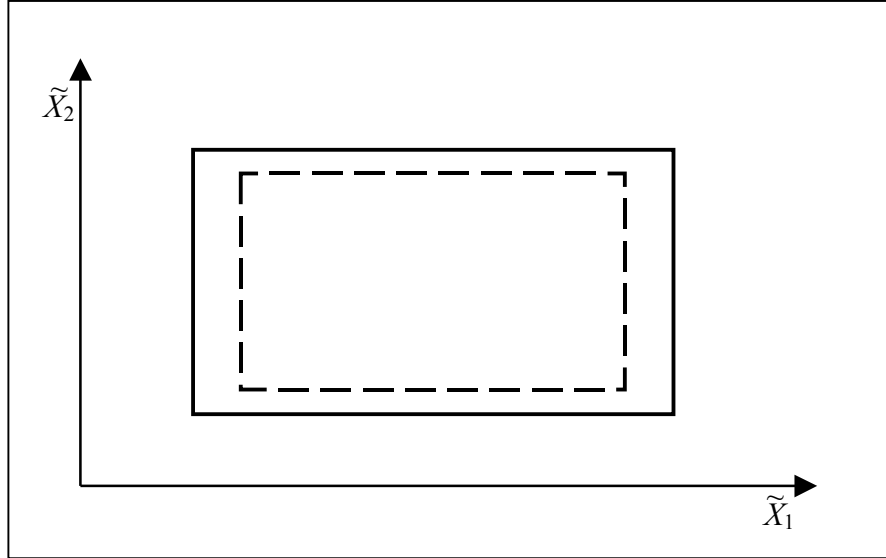
Starting from the estimated fuzzy data  $\tilde{x}_{ij}^*$ ,  $i=1,\dots,I$ ,  $j=1,\dots,J$ , we then get the estimated  $J$ -dimensional fuzzy vectors  $\tilde{\mathbf{x}}_i^* = (\tilde{x}_{i1}^*, \dots, \tilde{x}_{iJ}^*)$ ,  $i=1,\dots,I$ , which, under the non-interactivity assumption, have the following membership functions:

$$\mu_{\tilde{\mathbf{x}}_i^*}(\mathbf{u}) = \min_{j=1,\dots,J} \mu_{\tilde{x}_{ij}^*}(u_{ij}), \quad i=1,\dots,I. \quad (14)$$

In fact, our model consists of finding  $\tilde{\mathbf{A}}$  and  $\mathbf{F}$  such that the observed data are within the estimated ones at the level  $h$  (see Figure 2):

$$[\tilde{\mathbf{x}}_i]_h \subseteq [\tilde{\mathbf{x}}_i^*]_h, \quad i=1,\dots,I. \quad (15)$$

**Figure 2:**  $h$ -level set ( $0 < h < 1$ ) of non-interactive fuzzy numbers ( $J = 2$ ): inclusion relations between observed data (dotted line) and estimated data (continuous line).



Therefore, in the possibilistic regression framework, each observation unit is represented by a score on a single fuzzy variable (the output variable) which can be represented as a fuzzy interval in  $\mathfrak{R}^1$  and, thus, the inclusion relations at the level  $h$  refer to such an interval. On the contrary, in the possibilistic data reduction framework, the inclusion relations at the level  $h$  between the observed data and the estimated ones

are obtained considering all the  $K$  vertices which characterize each hyperrectangle pertaining to the  $J$ -dimensional observed and estimated fuzzy vectors (see Figure 2).

From (15), taking into account (10) and (12), we then have the following inclusion constraints:

$$\begin{aligned} m_{ij} + s_{ij}q_{kj} \left| L_{ij}^{-1}(h) \right| &\leq \mathbf{a}_{Mi} \mathbf{f}'_j + \mathbf{a}_{Si} \left| \mathbf{f}'_j \right| q_{kj} \left| L_{ij}^{-1}(h) \right| \quad \text{if } q_{kj} = 1, i=1, \dots, I; j=1, \dots, J; k=1, \dots, K; \\ m_{ij} + s_{ij}q_{kj} \left| L_{ij}^{-1}(h) \right| &\geq \mathbf{a}_{Mi} \mathbf{f}'_j + \mathbf{a}_{Si} \left| \mathbf{f}'_j \right| q_{kj} \left| L_{ij}^{-1}(h) \right| \quad \text{if } q_{kj} = -1, i=1, \dots, I; j=1, \dots, J; k=1, \dots, K. \end{aligned} \quad (16)$$

In (16),  $m_{ij}$  and  $s_{ij}$  are the generic elements of, respectively,  $\mathbf{M}$  and  $\mathbf{S}$  and the  $q_{kj}$ 's,  $j=1, \dots, J$ ;  $k=1, \dots, K$ ; help us to define every vertex of the hyperrectangle associated to each observation unit separately. In fact, they are the elements of the matrix  $\mathbf{Q}$  of order  $(K \times J)$  whose elements are equal to  $\pm 1$ . More specifically each row refers exactly to every vertex of the hyperrectangle by considering all the possible combinations of +1's and -1's. For example, if  $J = 2$ , we get:

$$\mathbf{Q} = \begin{bmatrix} -1 & -1 \\ -1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (17)$$

With respect to the generic  $i$ -th observation unit, we deal with the vector of the lower bounds by considering  $\mathbf{q}_1$ , the first row of  $\mathbf{Q}$ . In fact, it yields

$$\mathbf{m}_i + \mathbf{s}_i * \mathbf{q}_1 = (m_{i1} \quad m_{i2}) + (s_{i1} \quad s_{i2}) * (-1 \quad -1) = (m_{i1} \quad m_{i2}) - (s_{i1} \quad s_{i2}) = \mathbf{m}_i - \mathbf{s}_i. \quad (18)$$

where the symbol  $*$  denotes the Hadamard product, that is the elementwise product of two matrices (vectors) of the same order<sup>3</sup>. Instead, the vector of the upper bounds is obtained by considering  $\mathbf{q}_3$ , the third row of  $\mathbf{Q}$ :

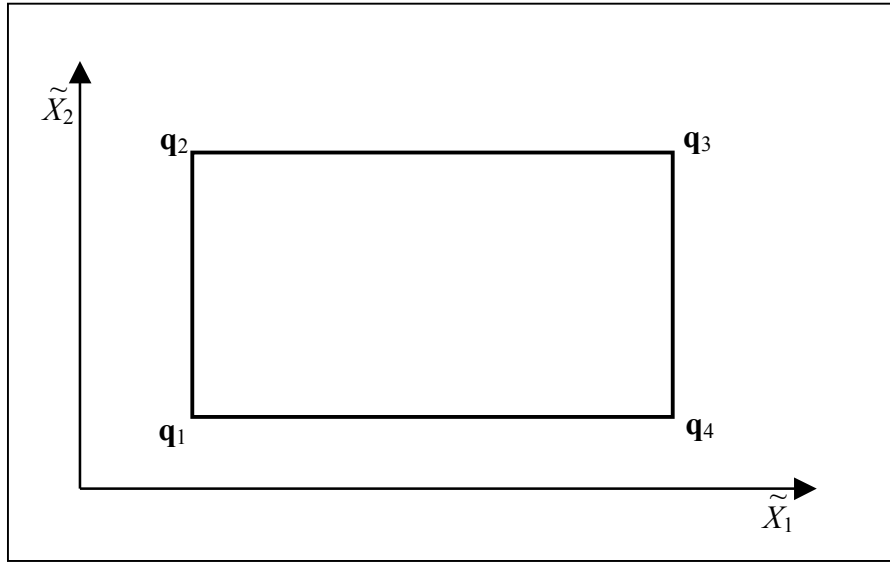
---

<sup>3</sup> Given two matrices  $\mathbf{Y}$  and  $\mathbf{Z}$  of order  $(u \times v)$ , we have  $\mathbf{Y} * \mathbf{Z} = \begin{bmatrix} y_{11}z_{11} & \cdots & y_{1v}z_{1v} \\ \vdots & \ddots & \vdots \\ y_{u1}z_{u1} & \cdots & y_{uv}z_{uv} \end{bmatrix}$

$$\mathbf{m}_i + \mathbf{s}_i * \mathbf{q}_3 = (m_{i1} \ m_{i2}) + (s_{i1} \ s_{i2}) * (1 \ 1) = (m_{i1} \ m_{i2}) + (s_{i1} \ s_{i2}) = \mathbf{m}_i + \mathbf{s}_i. \quad (19)$$

The remaining two vertices can be analogously obtained considering the remaining rows of  $\mathbf{Q}$  (see Figure 3). Also see, for more details about the matrix  $\mathbf{Q}$ , Giordani and Kiers (2004).

**Figure 3:** Strong  $h$ -level set of non-interactive fuzzy numbers ( $J = 2$ ) when  $h = 0$ .



**Proposition 1:** The inclusion relations among  $J$ -dimensional (observed and estimated) fuzzy numbers given in (16) hold when the inclusions among all the elements of the (observed and estimated)  $J$ -dimensional fuzzy vectors hold.

**Proof:** Let  $\mathbf{Q}_k$ 's be the diagonal matrices whose main diagonal elements are equal to those of  $\mathbf{q}_k$ , the  $k$ -th row of  $\mathbf{Q}$ ,  $k = 1, \dots, K$ . The following relations hold:

$$\mathbf{M} + (\mathbf{S}\mathbf{Q}_1) * \mathbf{L}_h = \mathbf{M} + \left( \mathbf{S} \begin{bmatrix} -1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & -1 \end{bmatrix} \right) * \mathbf{L}_h = \mathbf{M} - \mathbf{S} * \mathbf{L}_h \leq \mathbf{M} + (\mathbf{S}\mathbf{Q}_k) * \mathbf{L}_h, \quad k = 1, \dots, K; \quad (20)$$

$$\mathbf{M} + (\mathbf{S}\mathbf{Q}_{2^{J-1}+1}) * \mathbf{L}_h = \mathbf{M} + \left( \mathbf{S} \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & 1 \end{bmatrix} \right) * \mathbf{L}_h = \mathbf{M} + \mathbf{S} * \mathbf{L}_h \geq \mathbf{M} + (\mathbf{S}\mathbf{Q}_k) * \mathbf{L}_h, \quad k = 1, \dots, K; \quad (21)$$

where the generic element of  $\mathbf{L}_h$  is  $|L_{ij}^{-1}(h)|$ . Thus, the elements of the lower bound matrix (from (17), if  $J = 2$ , considering  $\mathbf{q}_1$ ) are lower than the elements of the matrices pertaining to the remaining vertices. Similarly, the elements of the upper bound matrix (from (17), if  $J = 2$ , considering  $\mathbf{q}_3$ ) are higher than those pertaining to the remaining vertices. Therefore, the inclusion constraints in (16) hold if the matrices  $\tilde{\mathbf{A}}$  and  $\mathbf{F}$  are such that the observed lower bounds are higher than the estimated ones and the observed upper bounds are lower than the estimated ones. In fact, the following inequalities hold:

$$\mathbf{M}^* - \mathbf{S}^* * \mathbf{L}_h \leq \mathbf{M} - \mathbf{S} * \mathbf{L}_h \leq \mathbf{M} + (\mathbf{S}\mathbf{Q}_k) * \mathbf{L}_h, \quad k = 1, \dots, K; \quad (22)$$

$$\mathbf{M}^* + \mathbf{S}^* * \mathbf{L}_h \geq \mathbf{M} + \mathbf{S} * \mathbf{L}_h \geq \mathbf{M} + (\mathbf{S}\mathbf{Q}_k) * \mathbf{L}_h, \quad k = 1, \dots, K. \quad (23)$$

Thus, for each observation unit and each variable, starting from  $K$  inclusion constraints, as given in (16), we reduce considering only two inclusion constraints. More specifically, for each pair  $(i, j)$ , we must only consider the inclusions between the observed and estimated lower and upper bounds, respectively:

$$\begin{aligned} m_{ij} - s_{ij} |L_{ij}^{-1}(h)| &\leq \mathbf{a}_{Mi} \mathbf{f}'_j - \mathbf{a}_{Si} |\mathbf{f}'_j| |L_{ij}^{-1}(h)| \quad i = 1, \dots, I; j = 1, \dots, J; \\ m_{ij} + s_{ij} |L_{ij}^{-1}(h)| &\geq \mathbf{a}_{Mi} \mathbf{f}'_j + \mathbf{a}_{Si} |\mathbf{f}'_j| |L_{ij}^{-1}(h)| \quad i = 1, \dots, I; j = 1, \dots, J. \end{aligned} \quad (24)$$

Now, from (10), taking into account (22) and (23), the constraints in (16) can be simplified as

$$\begin{aligned} \mathbf{M} + \mathbf{S} * \mathbf{L}_h &\leq \mathbf{A}_M \mathbf{F}' + (\mathbf{A}_S |\mathbf{F}'|) * \mathbf{L}_h, \\ \mathbf{M} - \mathbf{S} * \mathbf{L}_h &\geq \mathbf{A}_M \mathbf{F}' - (\mathbf{A}_S |\mathbf{F}'|) * \mathbf{L}_h. \end{aligned} \quad (25)$$

It is easy to see that (25) can be rewritten as

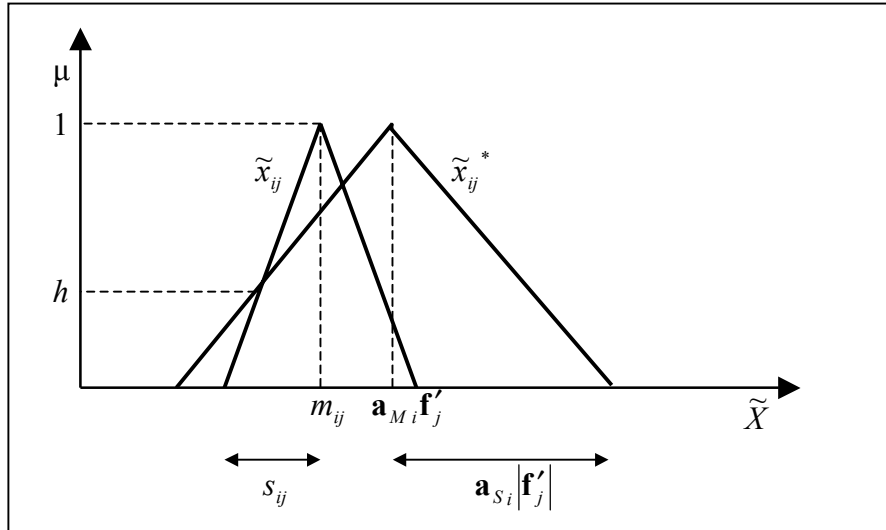
$$\begin{aligned}
 & \begin{bmatrix} m_{11} + s_{11} |L_{11}^{-1}(h)| & \cdots & m_{1J} + s_{1J} |L_{1J}^{-1}(h)| \\ \vdots & \ddots & \vdots \\ m_{I1} + s_{I1} |L_{I1}^{-1}(h)| & \cdots & m_{IJ} + s_{IJ} |L_{IJ}^{-1}(h)| \end{bmatrix} \leq \begin{bmatrix} \mathbf{a}_{M1} \mathbf{f}'_1 + \mathbf{a}_{S1} |\mathbf{f}'_1| |L_{11}^{-1}(h)| & \cdots & \mathbf{a}_{M1} \mathbf{f}'_J + \mathbf{a}_{SJ} |\mathbf{f}'_J| |L_{1J}^{-1}(h)| \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{MI} \mathbf{f}'_1 + \mathbf{a}_{SI} |\mathbf{f}'_1| |L_{I1}^{-1}(h)| & \cdots & \mathbf{a}_{MI} \mathbf{f}'_J + \mathbf{a}_{SI} |\mathbf{f}'_J| |L_{IJ}^{-1}(h)| \end{bmatrix} \\
 & \begin{bmatrix} m_{11} - s_{11} |L_{11}^{-1}(h)| & \cdots & m_{1J} - s_{1J} |L_{1J}^{-1}(h)| \\ \vdots & \ddots & \vdots \\ m_{I1} - s_{I1} |L_{I1}^{-1}(h)| & \cdots & m_{IJ} - s_{IJ} |L_{IJ}^{-1}(h)| \end{bmatrix} \geq \begin{bmatrix} \mathbf{a}_{M1} \mathbf{f}'_1 - \mathbf{a}_{S1} |\mathbf{f}'_1| |L_{11}^{-1}(h)| & \cdots & \mathbf{a}_{M1} \mathbf{f}'_J - \mathbf{a}_{SJ} |\mathbf{f}'_J| |L_{1J}^{-1}(h)| \\ \vdots & \ddots & \vdots \\ \mathbf{a}_{MI} \mathbf{f}'_1 - \mathbf{a}_{SI} |\mathbf{f}'_1| |L_{I1}^{-1}(h)| & \cdots & \mathbf{a}_{MI} \mathbf{f}'_J - \mathbf{a}_{SI} |\mathbf{f}'_J| |L_{IJ}^{-1}(h)| \end{bmatrix}. \quad (26)
 \end{aligned}$$

Thus, taking into account (11), (25) consists of imposing that all the inclusion relations among the  $IJ$  observed fuzzy data  $\tilde{x}_{ij}$ 's and estimated fuzzy data  $\tilde{x}_{ij}^*$ 's hold,  $i=1, \dots, I$ ,  $j=1, \dots, J$  (see Figure 4):

$$[\tilde{x}_{ij}]_h \subseteq [\tilde{x}_{ij}^*]_h, \quad i=1, \dots, I, \quad j=1, \dots, J. \quad (27)$$

□

**Figure 4:** Degree of fitting of  $\tilde{x}_{ij}^*$  to  $\tilde{x}_{ij}$ .



### 3.3. The minimization problem

Among all the feasible values of  $\tilde{\mathbf{A}}$  and  $\mathbf{F}$  for which the inclusion relations in (15) hold, we seek the optimal parameters in such a way that the sum of the estimated spreads (the



fuzziness) is minimized. From (9), we thus have:

$$\min J = \sum_{i=1}^I \sum_{j=1}^J \mathbf{a}_{Si} |\mathbf{f}'_j|. \quad (28)$$

As we already saw, the inclusion relation in (15) can be simplified by considering (25). It is worth to notice that in (25) as well as in (16) the elements of  $\mathbf{A}_S$  are constrained to be non-negative in order to avoid the possible negativity of the estimated spreads. In fact, we set

$$\mathbf{A}_S \geq \Psi. \quad (29)$$

where  $\Psi$  is a crisp matrix of the same order of  $\mathbf{A}_S$ , whose elements are non-negative. A recognized problem in possibilistic regression analysis (see, for instance, Tanaka and Lee, 1998) that, as we observed, also occurs in our possibilistic latent structure model is the tendency of some coefficients to become crisp because of the minimization problem. This problem can be solved by constraining the component score spreads to be higher than a pre-specified threshold, say  $\Psi$ .

As we already pointed out, we also constrain the matrix  $\mathbf{F}$  to be columnwise orthonormal:

$$\mathbf{F}'\mathbf{F} = \mathbf{I}. \quad (30)$$

This ensures the non-interactivity of the coefficients in  $\tilde{\mathbf{A}}$  and offers the opportunity to adequately represent the observation units as low-dimensional hyperrectangles (with edges parallel to the new axes), in the subspace spanned by  $\mathbf{F}$ , using the elements of  $\tilde{\mathbf{A}}$  as coordinates. Notice that, if  $\Psi=0$ , in the low dimensional representation,  $P$ -hyperrectangles having  $P$ -dimensional volume equal to 0 may occur because of at least one component score spread being equal to 0. It is important to observe that, the estimated centers ( $\mathbf{A}_M \mathbf{F}'$ ) give the orthogonal projection of  $\mathbf{M}$  onto the subspace spanned by  $\mathbf{F}$  using  $\mathbf{A}_M \mathbf{F}'$  as coordinates. Such a property does not hold for the spreads

matrix because of the absolute value of  $\mathbf{F}$  in (25) and (28) which helps to avoid that negative estimates of the spreads occur. Thus, the size of the low dimensional hyperrectangles only provides a measure of the vagueness associated with each observation unit. Future research is needed to improve the visualization tool.

Taking into account (25), (28), (29) and (30), the latent structure problem can be formulated in terms of the following classical non-linear programming (NLP) problem:

$$\begin{aligned}
& \min \sum_{i=1}^I \sum_{j=1}^J \mathbf{a}_{S_i} |\mathbf{f}'_j| \\
& \text{s.t. } \mathbf{M} + \mathbf{S} * \mathbf{L}_h \leq \mathbf{A}_M \mathbf{F}' + (\mathbf{A}_S |\mathbf{F}'|) * \mathbf{L}_h, \\
& \quad \mathbf{M} - \mathbf{S} * \mathbf{L}_h \geq \mathbf{A}_M \mathbf{F}' - (\mathbf{A}_S |\mathbf{F}'|) * \mathbf{L}_h, \\
& \quad \mathbf{F}' \mathbf{F} = \mathbf{I}, \\
& \quad \mathbf{A}_S \geq \mathbf{\Psi}.
\end{aligned} \tag{31}$$

**Proposition 2:** There always exists a solution of the NLP problem in (31).

**Proof:** Given a feasible solution for  $\mathbf{F}$  such that the constraint in (30) is satisfied, we can find a feasible solution for  $\tilde{\mathbf{A}}$  according to the constraints in (25) by taking a sufficiently large positive matrix for  $\mathbf{A}_S$ .

□

**Remark 1 (Triangular membership function)**

When a triangular fuzzy number is used,  $L(x) = \max(0, 1 - |x|)$ . It follows that  $|L^{-1}(h)| = 1 - h$ .

**Remark 2 (Number of latent variables)**

In order to detect the optimal number of latent variables, we suggest to choose  $P$  such that it can be considered optimal in performing the classical latent structure model, as described in Section 2, on the (crisp) centers matrix.

**Remark 3 (Preprocessing)**

Prior to applying the latent structure analysis, it could be recommended to preprocess the data in order to eliminate unwanted differences in level and scale among variables.

To do so, following Giordani and Kiers (2004), we suggest to standardize each center by subtracting the mean of the centers of the variable at hand and by dividing each center by the standard deviation of the centers. Finally, we suggest to divide the spreads by the standard deviation of the corresponding center.

**Remark 4 (The special case  $P = J$  latent variables)**

When the number of latent variables equals that of observed ones, the minimization problem in (31) has the solution

$$\begin{aligned}\mathbf{A}_M &= \mathbf{M}, \\ \mathbf{A}_S &= \mathbf{S}, \\ \mathbf{F} &= \mathbf{I},\end{aligned}\tag{32}$$

and the following equalities hold:

$$\begin{aligned}\mathbf{M} + \mathbf{S} * \mathbf{L}_h &= \mathbf{A}_M \mathbf{F}' + (\mathbf{A}_S |\mathbf{F}'|) * \mathbf{L}_h = \mathbf{M}^* + \mathbf{S}^* * \mathbf{L}_h, \\ \mathbf{M} - \mathbf{S} * \mathbf{L}_h &= \mathbf{A}_M \mathbf{F}' - (\mathbf{A}_S |\mathbf{F}'|) * \mathbf{L}_h = \mathbf{M}^* + \mathbf{S}^* * \mathbf{L}_h.\end{aligned}\tag{33}$$

Thus, the estimated and observed data coincide.

**Remark 5 (Interval Possibilistic Latent Structure Model)**

The described problem deals with symmetric fuzzy data. The special case when the data are intervals can also be proposed. Following the regression framework, the latent variables are determined such that the estimated intervals totally enclose the observed interval valued data. Thus, (25) is replaced by

$$\begin{aligned}\mathbf{M} + \mathbf{S} &\leq \mathbf{A}_M \mathbf{F}' + \mathbf{A}_S |\mathbf{F}'|, \\ \mathbf{M} - \mathbf{S} &\geq \mathbf{A}_M \mathbf{F}' - \mathbf{A}_S |\mathbf{F}'|,\end{aligned}\tag{34}$$

where  $\mathbf{S}$  is now the radii matrix (the radius is the half-width of an interval) and  $\mathbf{A} = (\mathbf{A}_M, \mathbf{A}_S)$  is the interval valued component score matrix. According to (26) the

intervals are obtained in such a way that the uncertainty is minimized. Therefore, considering (28), (29), (30) and (34) yields the following NLP problem:

$$\begin{aligned}
& \min \sum_{i=1}^I \sum_{j=1}^J \mathbf{a}_{Si} |\mathbf{f}'_j| \\
& \text{s.t. } \mathbf{M} + \mathbf{S} \leq \mathbf{A}_M \mathbf{F}' + (\mathbf{A}_S |\mathbf{F}'|), \\
& \quad \mathbf{M} - \mathbf{S} \geq \mathbf{A}_M \mathbf{F}' - (\mathbf{A}_S |\mathbf{F}'|), \\
& \quad \mathbf{F}' \mathbf{F} = \mathbf{I}, \\
& \quad \mathbf{A}_S \geq \mathbf{\Psi}.
\end{aligned} \tag{35}$$

**Remark 6 (Lower problem)**

In the minimization problem in (31), the fuzziness is minimized as much as possible in such a way that the constraints in (15) hold. Following the possibilistic regression framework, we refer to the minimization problem in (31) as *Upper Model*. Analogously, we propose the *Lower Model* where the problem is to satisfy

$$[\tilde{\mathbf{x}}_i]_h \supseteq [\tilde{\mathbf{x}}_i^*]_h, \quad i=1, \dots, I, \tag{36}$$

and to maximize the fuzziness (the sum of the spreads). Thus, whereas in the *Upper Model*, we aim at finding the smallest fuzzy set which satisfies (15), in the *Lower Model*, we aim at finding the largest fuzzy set which satisfies (36). In this case, the solution is obtained by solving the following NLP problem:

$$\begin{aligned}
& \max \sum_{i=1}^I \sum_{j=1}^J \mathbf{a}_{Si} |\mathbf{f}'_j| \\
& \text{s.t. } \mathbf{M} + \mathbf{S} * \mathbf{L}_h \geq \mathbf{A}_M \mathbf{F}' + (\mathbf{A}_S |\mathbf{F}'|) * \mathbf{L}_h, \\
& \quad \mathbf{M} - \mathbf{S} * \mathbf{L}_h \leq \mathbf{A}_M \mathbf{F}' - (\mathbf{A}_S |\mathbf{F}'|) * \mathbf{L}_h, \\
& \quad \mathbf{F}' \mathbf{F} = \mathbf{I}, \\
& \quad \mathbf{A}_S \geq \mathbf{\Psi}.
\end{aligned} \tag{37}$$

We notice that the problem in (37) may not have any optimal solution, as it happens in the regression framework. In fact, even taking  $\mathbf{A}_S = \mathbf{0}$  does not guarantee that a feasible solution is obtained.

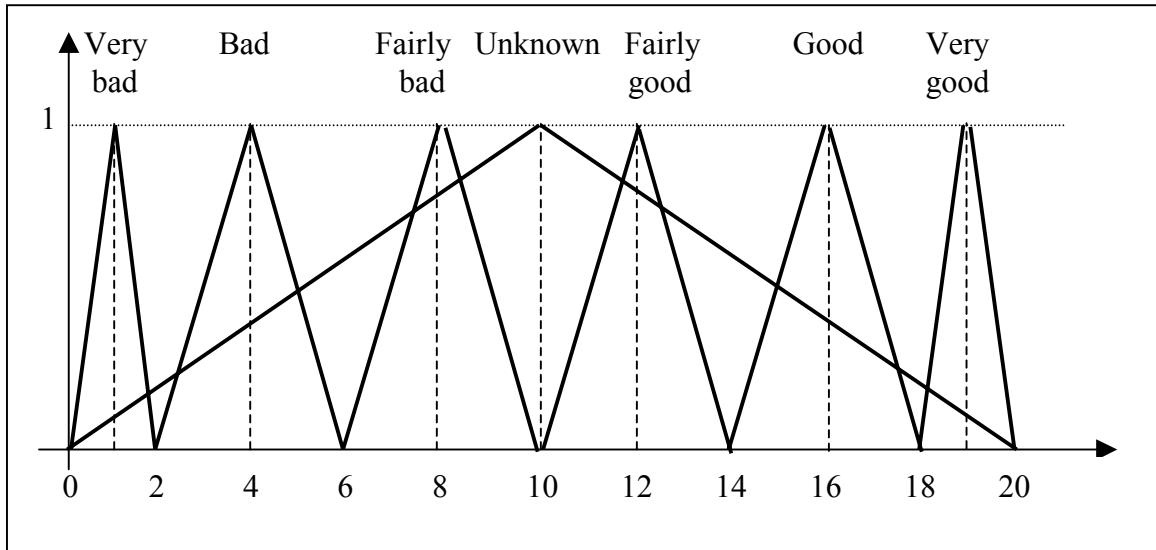
## 4. Applications

In this section we provide two applications of our method in order to show how it works.

### 4.1. Student data

This section is devoted to the application of the here-proposed model to the Student data set given in Table 1. The data refer to  $J = 4$  marks obtained by  $I = 6$  students in mathematics and physics. The marks take values from 0 to 20. The data are fuzzy because one mark is unknown, several mark are imprecise and are given by intervals or linguistic labels. We consider plausible to assume that each mark is a symmetric fuzzy number with a triangular membership function. The linguistic labels are fuzzified according to Figure 5.

**Figure 5:** Fuzzification of the linguistic labels in the Student data set



The data are preprocessed by centering the centers. After running several analyses, we decide to set  $P = 2$  and  $h = 0$ . Moreover, to avoid that  $\mathbf{A}_s$  has zero elements, we set  $a_{s31} = 0.10$ . The increase of the minimization function is negligible (the loss function value is only 0.01% higher than that using  $\Psi=0$ ).

We obtain the fuzzy component score and (crisp) component loading matrices, respectively, in Table 2 and Table 3.

**Table 2:** Fuzzy component score matrix (*center, spread*)<sub>L</sub>

Student	PC1	PC2
Tom	(5.67, 10.31) <sub>L</sub>	(-2.75, 15.39) <sub>L</sub>
David	(5.04, 5.22) <sub>L</sub>	(-1.25, 2.48) <sub>L</sub>
Bob	(-2.11, 0.10) <sub>L</sub>	(4.52, 4.72) <sub>L</sub>
Jane	(6.99, 3.44) <sub>L</sub>	(3.30, 9.82) <sub>L</sub>
Joe	(-6.76, 5.84) <sub>L</sub>	(-0.74, 4.47) <sub>L</sub>
Jack	(-8.32, 0.99) <sub>L</sub>	(-7.54, 3.29) <sub>L</sub>

**Table 3:** Crisp component loading matrix

Mark	PC1	PC2
Mathematics 1	0.70	0.02
Mathematics 2	0.71	0.09
Physics 1	-0.03	0.74
Physics 2	-0.08	0.67

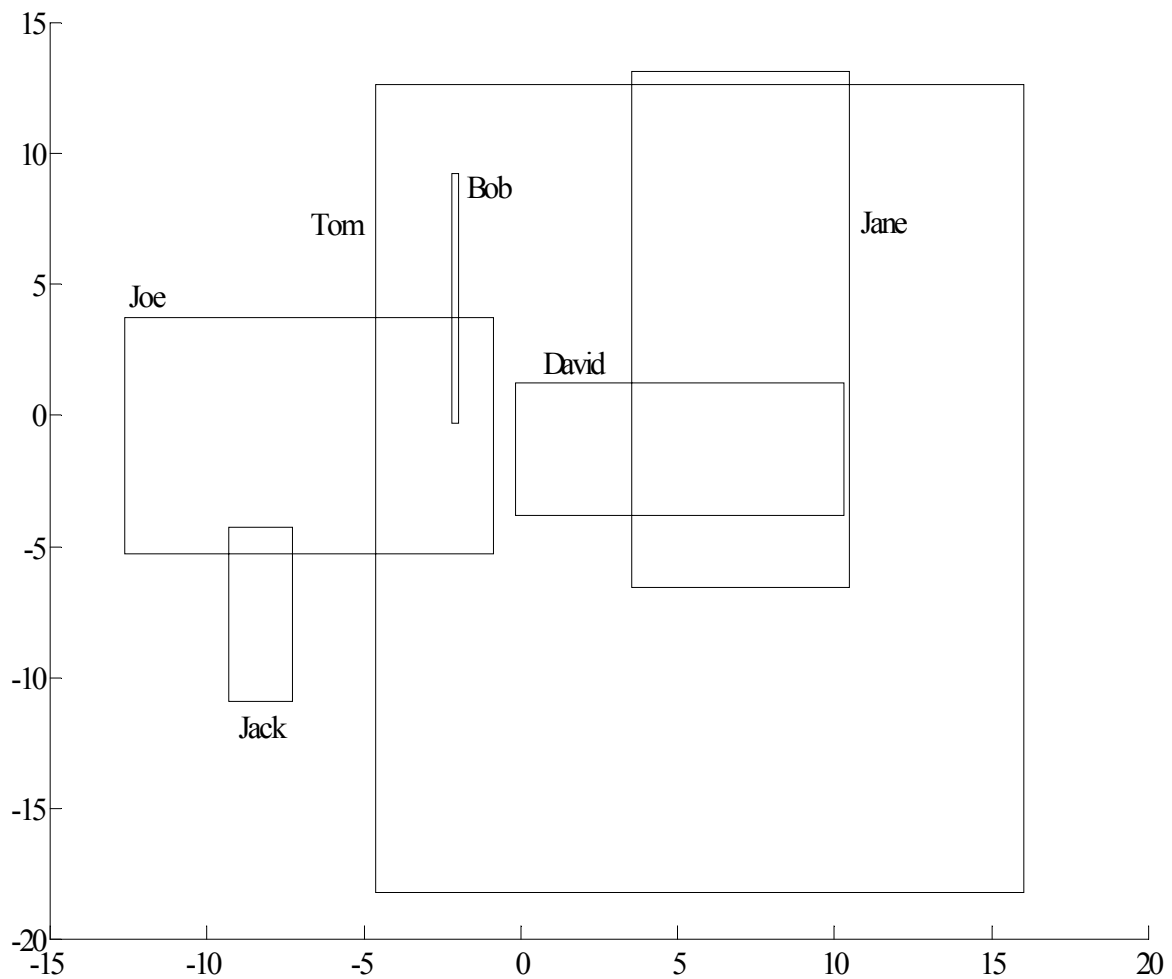
From Table 3, we can easily interpret the latent variables. The first one has high loadings for Mathematics 1 and 2. The remaining loadings take low values. Thus, such a latent variable reflects the mathematical knowledge of the students. On the contrary, the second latent variable is the physical knowledge. In fact, high loadings pertain to Physics 1 and 2 whereas the remaining loadings are negligible.

By considering the fuzzy component score matrix, we can assess the ability of the students in mathematics and physics. Specifically, the centers provide information about the knowledge of the students in mathematics (first component) and physics (second component). The spreads give a measure of the uncertainty associated to each center. Therefore, the fuzzy component scores are the (fuzzy) coefficients that relate the observed variables to the latent ones. Specifically, high scores denote high levels of knowledge in the involved matter. In fact, we can observe that Jane has the highest first component score. Thus, Jane is the best student according to the mathematical marks. The following rating can be found with respect to the mathematical knowledge: Jane, Tom, David, Bob, Joe, Jack. This result is consistent with the observed marks given in Table 1. Similar results can be obtained considering the second component. Now, Bob

and, then, Jane are the best students in physics. Then, Joe, David, Tom and Jack follow. Thus, Jack has the lowest component scores with respect to either component. Therefore, the analysis shows that he is the worst student as one may also observe from Table 1. The uncertainty of the component scores (the size of the spreads) is related to the uncertainty of the observed data. For instance, notice that Tom is characterized by the highest spread component scores. This depends on the fact that one mark (P1) is unknown. It follows that such an observed fuzziness implies extremely fuzzy component scores.

Further details can be found in Figure 6 where we plot the students (as rectangles) in the obtained low dimensional space.

**Figure 6:** Low dimensional configuration of the students



As one may expect, the rectangle pertaining to Tom is the biggest one. In fact, the size of the rectangles reflects the fuzziness of the observed data. It is interesting to observe the rectangle pertaining to Bob. It is characterized by a very small basis and a high height. This can be explained by considering the interpretation of the component. The first component depends on the marks in mathematics. By observing Table 1, the fuzziness of the Bob's marks in mathematics is very low (the marks are 6 and [10,11]). On the contrary, the fuzziness of the marks in physics ([13,20] and 'good') is captured by the height of the involved rectangle.

Taking into account the interpretation of the latent variables, the best students are on the upper right side whereas the worst ones are on the lower left side of the Figure. High levels of knowledge in mathematics and low levels of knowledge in physics characterize the students located on the upper left side of Figure 6. The opposite comments hold considering the students located on the lower right side of the figure. Thus, as Jane is located in the upper right side of Figure 6, she seems to be the best student according to the marks in mathematics and physics. Bob and David are, respectively, on the high side and on the upper side of the figure. In fact, Bob has remarkable results in physics and David has quite good results in mathematics. Tom has high marks in both mathematics and physics. Notice that the rectangle pertaining to Tom is on the right side of the figure. As one mark in physics (P1) pertaining to Tom is unknown, the size and the position of the rectangle with respect to the second component reflect such an uncertainty. In fact, the second component depends on the marks in physics and, therefore, the involved rectangle does not provide information in order to evaluate the Tom's knowledge in physics. This should be seen as a nice property of the models because the obtained low dimensional configuration of the observation units does not offer misleading results that can arise by simply considering the centers information. Finally, inspecting Figure 6 again, Joe and, above all, Jack (whose rectangle is located on the left lower side of the figure) seem to be the worst students.



#### 4.2. Fats and oils data

The data set involved is the well known “Fats and oils data” (Ichino, 1988) which describes  $I=8$  fats and oils (two animal fats and six vegetable oils) by means of  $J=4$  interval valued variables. The original data set has one additional qualitative variable that is not taken into account in the current application. The data are summarized in Table 5. The variables are ‘Specific Gravity’, ‘Freezing Point’, ‘Iodine Value’ and ‘Saponification’.

**Table 4:** Fats and oils data

Fats and Oils	Specific Gravity	Freezing Point	Iodine Value	Saponification
Linseed oil	[0.930,0.935]	[-27,-18]	[170,204]	[118,196]
Perilla oil	[0.930,0.937]	[-5,-4]	[192,208]	[188,197]
Cottonseed oil	[0.916,0.918]	[-6,-1]	[99,113]	[189,198]
Sesame oil	[0.920,0.926]	[-6,-4]	[104,116]	[187,193]
Camellia oil	[0.916,0.917]	[-21,-15]	[80,82]	[189,193]
Olive oil	[0.914,0.919]	[0,6]	[79,90]	[187,196]
Beef tallow	[0.860,0.870]	[30,38]	[40,48]	[190,199]
Hog fat	[0.858,0.864]	[22,32]	[53,77]	[190,202]

Note: the first number is the lower bound, the latter one is the upper bound.

After preprocessing the data according to Remark 3, we perform the FP-PCA model. More specifically, we solve the NLP problem for interval valued data in (35) setting  $\Psi_{12}=0.05$  and the remaining elements of  $\Psi$  equal to 0. We set only one lower bound higher than 0 because, when  $\Psi=\mathbf{0}$ , the second component score spread pertaining to Linseed oil is 0 and, therefore, the low dimensional hyperrectangle involved is a segment. The increase of the minimization function is negligible (just 0.1% higher than that using  $\Psi=\mathbf{0}$ ).

We obtain the interval valued component score and (crisp) component loading matrices, respectively, in Table 5 and Table 6.

**Table 5:** Interval valued component score matrix

Fats and Oils	PC1	PC2
Linseed oil	(3.38,3.88)	(1.58,0.05)
Perilla oil	(1.16,1.78)	(1.01,0.55)
Cottonseed oil	(-0.46,0.42)	(0.41,0.16)
Sesame oil	(-0.08,0.27)	(0.70,0.27)
Camellia oil	(-0.60,0.60)	(1.08,0.64)
Olive oil	(-0.45,0.54)	(0.13,0.42)
Beef tallow	(-1.08,0.46)	(-2.44,0.27)
Hog fat	(-1.13,0.50)	(-2.16,0.62)

Note : the first number is the center, the latter one is the radii (half widths).

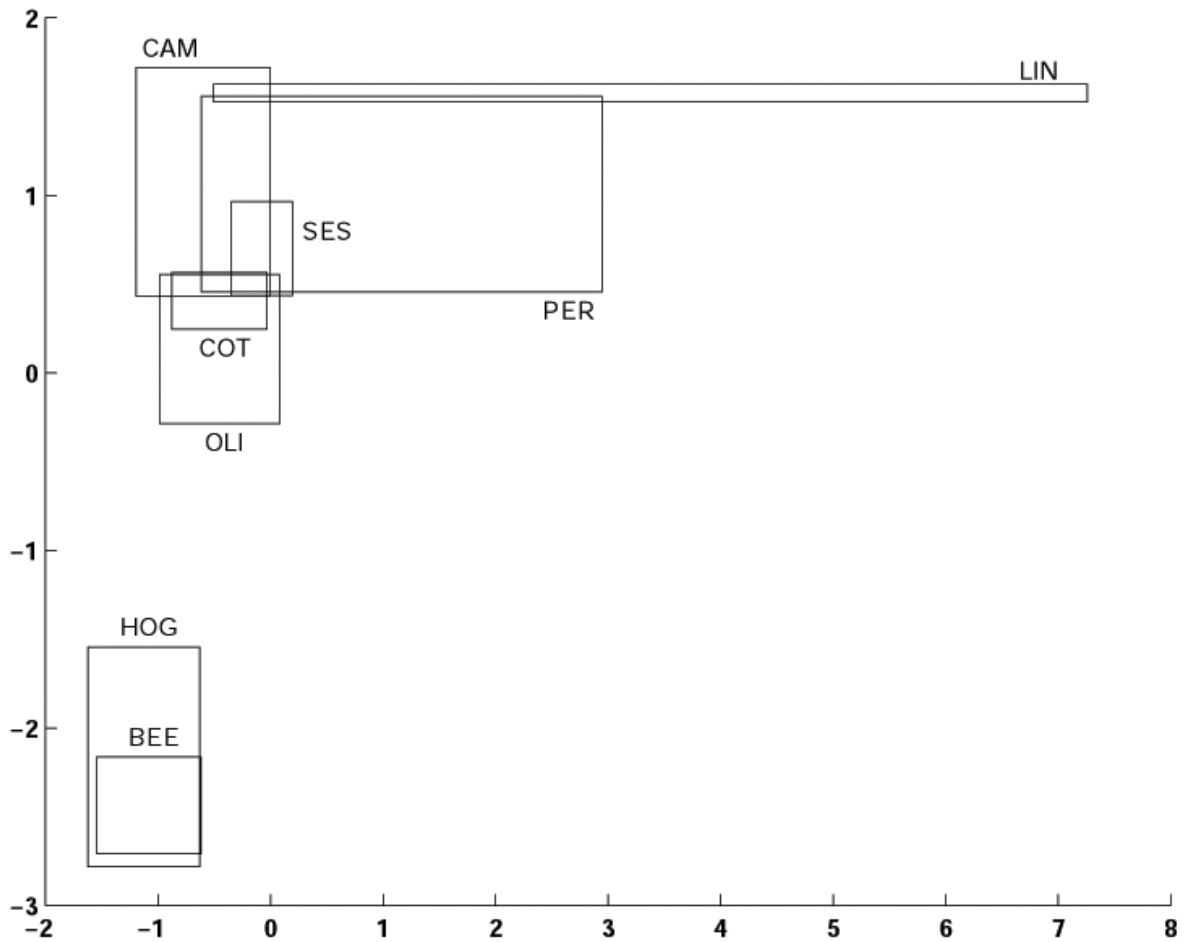
**Table 6:** Component loading matrix

Features	PC1	PC2
Specific Gravity	-0.03	0.68
Freezing Point	-0.05	-0.70
Iodine Value	0.54	0.18
Saponification	-0.84	0.14

By observing the component loading matrix, the components can be easily interpreted. More specifically, loading values far from 0 imply that the original variables involved play a relevant role in describing the extracted components. Thus, with respect to the first component, the scores are higher if the Iodine value is high and the Saponification is low. Specific gravity and Freezing point have a negligible role. In fact, they are relevant in defining the second component for which the loadings are, respectively, 0.68 and -0.70.

Taking into account that the component loading matrix is columnwise orthonormal, we can provide a low dimensional representation of the fats and oils as  $P$ -dimensional hyperrectangles in the space spanned by the columns of  $\mathbf{F}$ . The coordinates of the observation units are given by the rows of the component score matrix. We get the plot in Figure 7.

**Figure 7:** Low dimensional representation of the fats and oils



From Figure 7, we can observe that the oils on the right (Linseed oil and Perilla oil) have the highest Iodine values according to the interpretation of the first component. Linseed oil position also depends on the Saponification value which is the lowest among the fats and oils. On the left of Figure 7, Beef tallow and Hog fat are visible. Their positions are consistent with their low Iodine values and high Saponification values.

The low dimensional plot is also consistent with the second component loadings. In fact, the oils and fats are positioned in such a way that in the upper (lower) side of Figure 7 there are the observation units which are characterized by high (low) Specific gravity values and low (high) Freezing point values. Since the animal fats take low second component scores and the vegetal oils take high low second component scores. Thus, by inspecting Figure 7, we can conclude that the second latent variable can be interpreted as the type of oil (vegetal vs animal).

## 5. Conclusion

In many applications, when the data are crisp, it is common to compress and synthesize huge amounts of data. Unfortunately, it may happen that the empirical (observational or experimental) information and/or the system structure are fuzzy. For this reason, in this paper, we suggest an extension of latent structure models in order to study the data reduction problem for symmetric fuzzy data sets in a possibilistic environment. In particular, for fitting the fuzzy latent structure model, analogously to possibilistic regression (Tanaka et al., 1982), we use a minimum fuzziness criterion.

The classical PCA for crisp data given in (1) involves the computation of the eigenvalues and eigenvectors of the (crisp) cross-product matrix  $\mathbf{X}'\mathbf{X}$ <sup>4</sup>. The solution is then found by considering the first  $P$  (number of extracted components) highest eigenvalues. The optimal value of  $P$  is found by considering the amount of explained variance that increases when  $P$  increases. Notice also that the solutions are nested. Our extension, based on the fuzzy adjustment of (1) given in (8), does not exploit the (fuzzy) eigendecomposition of the (fuzzy) cross-product matrix but NLP tools.

The here proposed method must be considered as a starting point to extend latent structure models to deal with fuzzy data following a possibilistic approach. Much work remains to be done. In our way of thinking it can be improved along at least four directions. It is advisable to develop a tool for determining the optimal number of extracted components without utilizing information given by classical PCA. A nice property of PCA is that the solution is not unique. In fact, equally well fitting solutions can be obtained by arbitrary non-singular transformations of the score (loading) component matrix, provided that these rotations are compensated in the loading (score) component matrix. This helps in rotating the solution to simpler structure (with many loadings close to 0), that is to find a solution such that the interpretation is less intractable, as sometimes is the case. Our proposal leads to a unique solution. Thus, it does not allow to rotate the solution to simple structure. In order to avoid the tendency of some coefficients to become crisp because of the minimization problem, it can be

---

<sup>4</sup> Notice that, for any matrix  $\mathbf{Y}$  of order  $(u \times v)$  with  $u \geq v$ , by means of the SVD decomposition we have  $\mathbf{Y}=\mathbf{P}\mathbf{D}\mathbf{Q}'$  where  $\mathbf{P}$ , of order  $(u \times u)$ , and  $\mathbf{R}$ , of order  $(v \times v)$ , are matrices containing the unit length (left and right) singular vectors of  $\mathbf{Y}$  and  $\mathbf{D}$  is the diagonal matrix displaying on the diagonal the singular values of  $\mathbf{Y}$ . Notice that  $\mathbf{P}'\mathbf{P}=\mathbf{I}$  and  $\mathbf{Q}'\mathbf{Q}=\mathbf{I}$ . For any symmetric matrix  $\mathbf{Z}$  of order  $(u \times u)$ , by means of the eigendecomposition, we have  $\mathbf{Z}=\mathbf{K}\mathbf{U}\mathbf{K}'$ , where  $\mathbf{K}$  contains the unit length eigenvectors of  $\mathbf{Z}$  ( $\mathbf{K}'\mathbf{K}=\mathbf{I}$ ) and  $\mathbf{U}$  contains the eigenvalues of  $\mathbf{Z}$ . As  $\mathbf{Y}'\mathbf{Y}$  is a symmetric matrix of order  $(u \times u)$ , we have  $\mathbf{Y}'\mathbf{Y} = \mathbf{K}\mathbf{U}\mathbf{K}' = \mathbf{P}\mathbf{D}\mathbf{Q}'\mathbf{Q}\mathbf{D}\mathbf{P}' = \mathbf{P}\mathbf{D}'\mathbf{P}'$ . Thus, the singular values of  $\mathbf{Y}$  are the square root eigenvalues of  $\mathbf{Y}'\mathbf{Y}$ . The same result holds for  $\mathbf{Y}$  and  $\mathbf{Y}\mathbf{Y}'$ . See, for more details, ten Berge (1993).

advisable to consider a quadratic programming approach, instead of constraining the component score spreads to be higher than a pre-specified threshold. Finally, as noticed in section 3.3, the visualization tool should be improved.

Moreover, it will be interesting in future to set up possibilistic models for a general class of fuzzy data, i.e. LR fuzzy data, and formalize the analysis in a three-way framework, in order to study the data reduction problem for LR fuzzy three-way data array, e.g.,  $\tilde{\mathbf{X}} \equiv \{\tilde{x}_{ijt} = (m_{ijt}, s_{ijt}, v_{ijt})_{LR} : i = 1, \dots, I; j = 1, \dots, J; t = 1, \dots, T\}$ , where  $i, j$  and  $t$  denote the units, variables and occasions (times, spaces, etc.), respectively;  $\tilde{x}_{ijt} = (m_{ijt}, s_{ijt}, v_{ijt})_{LR}$  represents the LR fuzzy variable  $j$  observed on the  $i$ -th unit at occasion  $t$ , where  $m_{ijt}$  denotes the center and  $s_{ijt}$  and  $v_{ijt}$  the left and right spread, respectively, with the following membership function:

$$\mu(u_{ijt}) = \begin{cases} L\left(\frac{m_{ijt} - u_{ijt}}{s_{ijt}}\right) & u_{ijt} \leq m_{ijt} \quad (s_{ijt} > 0) \\ R\left(\frac{u_{ijt} - m_{ijt}}{v_{ijt}}\right) & u_{ijt} \geq m_{ijt} \quad (v_{ijt} > 0), \end{cases} \quad (38)$$

where  $L$  (and  $R$ ) is a decreasing “shape” function from  $\Re^+$  to  $[0,1]$  with  $L(0)=1$ ;  $L(z_{ijt}) < 1$  for all  $z_{ijt} > 0$ ,  $\forall i, j, t$ ;  $L(z_{ijt}) > 0$  for all  $z_{ijt} < 1$   $\forall i, j, t$ ;  $L(1)=0$  (or  $L(z_{ijt}) > 0$  for all  $z_{ijt}$  and  $L(+\infty)=0$ ).

Other extensions, regarding the theoretical structure of data reduction model, can be considered. For instance, it will be interesting to assume interrelations among fuzzy variables as in Celminš (1987) for the fuzzy least-squares regression. Moreover, analogously to the regression framework, we can measure the vagueness of the problem not only as the sum of the spreads, but in different ways, and assume additional hypothesis in order to reduce the constraints (Bardossy, 1990; Ruspini, et al., 1998).

## Acknowledgments

We would like to thank the Associate Editor and the Referees for their fruitful comments and suggestions.

## References

- Bardossy, A., Note on fuzzy regression, *Fuzzy Sets and Systems*, 37 (1990) 65-75.
- Barni, M., Cappellini, V., and Mecocci, A., Comments on 'A possibilistic approach to clustering', *IEEE Transactions on Fuzzy Systems*, 4 (1996) 393-396.
- Celminš, A., Least squares model fitting to fuzzy vector data, *Fuzzy Sets and Systems*, 22 (1987) 245-269.
- Celminš, A., A practical approach to nonlinear fuzzy regression, *SIAM Journal of Sci. Stat. Computing*, 12 (1991) 521-546.
- Chang, P.T., Lee, E.S., and Konz, S.A., Applying fuzzy linear regression to VDT legibility, *Fuzzy Sets and Systems*, 80 (1996) 197-204.
- Chang, P.T., Huang, L.C., and Lin, H.J., The fuzzy Delphi method via fuzzy statistics and membership function fitting and an application to the human resources, *Fuzzy Sets and Systems*, 112 (2000) 511-520.
- Chen, S.M., A new method for tool steel selection under fuzzy environment, *Fuzzy Sets and Systems*, 92 (1997) 265-274.
- Chen, Y.S., Outliers detection and confidence interval modification in fuzzy regression, *Fuzzy Sets and Systems*, 119 (2001) 259-272.
- Cheng, C.H., Yang, K.L., and Hwang, C.L., Evaluating attack helicopters by AHP based on linguistic variable weight, *European Journal of Operational Research*, 116 (1999) 423-435.
- Cheng, C.H., and Lin, Y., Evaluating the best main battle tank using fuzzy decision theory with linguistic criteria evaluation, *European Journal of Operational Research*, 142 (2002) 174-186.
- Chung, L.L.H., and Chan, K.C.C., Evolutionary discovery of fuzzy concepts in data, *Brain and Mind*, 4 (2003) 253-268.
- Coppi R., and D'Urso, P., Three-way fuzzy clustering models for LR fuzzy time trajectories, *Computational Statistics & Data Analysis*, 43 (2003) 149-177.
- Denceux, T., and Masson, M., Principal component analysis of fuzzy data using autoassociative neural networks, *IEEE Transactions on Fuzzy Systems* (2003) in press.
- Dubois, D., and Prade, H., *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.
- D'Urso, P., Linear regression analysis for fuzzy/crisp input and fuzzy/crisp output data, *Computational Statistics & Data Analysis*, 42 (1-2) (2003) 47-72.
- D'Urso, P., and Gastaldi, T., An "orderwise" polynomial regression procedure for fuzzy data, *Fuzzy Sets and Systems*, 130 (1) (2002) 1-19.
- Di Lascio, L., Gisolfi, A., Albunia, A., Galarti G., and Meschi, F., A fuzzy-based methodology for the analysis of diabetic neuropathy, *Fuzzy Sets and Systems*, 129 (2002), 203-228.
- Finol J., Guo, Y.K., and Jing, X.D., A rule based fuzzy model for the prediction of petrophysical rock parameters, *Journal of Petroleum Science & Engineering*, 29 (2001) 97-113.
- Giordani, P., and Kiers, H.A.L., Principal Component Analysis of symmetric fuzzy data, *Computational Statistics and Data Analysis* (2004) in press.
- Hamdi, M., Lachiver, G., and Michaud, F., A new predictive thermal sensation index of human response, *Energy and Buildings*, 29 (1999) 167-178.
- Herrera, F., Herrera-Viedma, E., and Verdegay, J.L., Linguistic measures based on fuzzy coincidence for researching consensus in group decision making, *International Journal of Approximate Reasoning*, 16 (1997) 309-334.
- Hisdal, E., Conditional possibilities, independence and noninteraction, *Fuzzy Sets and Systems*, 1 (1978) 283-297.
- Hougaard, J.L., Fuzzy scores of technical efficiency, *European Journal of Operational Research*, 115 (1999) 529-541.
- Huang, D., Chen, T., and Wang, M.J.J., A fuzzy set approach for event tree analysis, *Fuzzy Sets and Systems*, 118 (2001) 153-165.

- Ichino, M., 1988. General metrics for mixed features – the cartesian space theory for pattern recognition. In: Proc. 1988 IEEE Intern. Conf. on Systems, Man and Cybernetics, vol. 1, International Academic Publisher, Beijing and Shenyang, 494-497.
- Kallithraka, S., Arvanitoyannis, Kefalas, P., El-Zajouli, A., Soufleros, E., and Psarra, E., Instrumental and sensory analysis of Greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin, Food Chemistry, 73 (2001) 501-514.
- Kuncheva, L., and Steimann, F., Fuzzy diagnosis, Artificial Intelligence in Medicine, 16 (1999) 121-128.
- Krishnapuram, R., and Keller, J., The possibilistic *c*-means algorithm: insights and recommendations, IEEE Transactions on Fuzzy Systems, 4 (1996) 385-393.
- Laviolette, M., Seaman, J.W., Barrett, J.D., and Woodall, W.H., A probabilistic and statistical view of fuzzy methods, with discussion, Technometrics, 37 3 (1995) 249-292.
- Lee, H.M., Applying fuzzy set theory to evaluate the rate of aggregative risk in software development, Fuzzy Sets and Systems, 79 (1996) 323-336.
- Lee, H.S., Optimal consensus of fuzzy opinions under group decision making environment, Fuzzy Sets and Systems, 132 (2002) 303-315.
- Krishnapuram, R., and Keller, J., A possibilistic approach to clustering, IEEE Transactions on Fuzzy Systems, 1 (1993) 98-110.
- Lee, H.T., and Chen, S.H., Fuzzy regression model with fuzzy input and output data for manpower forecasting, Fuzzy Sets and Systems, 119 (2001) 205-213.
- Ménard, M., Courboulay, V., and Dardignac, P.A., Possibilistic and probabilistic fuzzy clustering: unification within the framework of the non-extensive thermostatics, Pattern Recognition, 36 (2003) 1325-1342.
- Moon, J.H., and Kang, C.S., Use of fuzzy set theory in the aggregation of expert judgments, Annals of Nuclear Energy, 26 (1999) 461-469.
- Nahmias, S., Fuzzy variables, Fuzzy Sets and Systems, 1 (1978) 97-110.
- Nakamori, Y., and Watada, J., Factor analysis for fuzzy data, in: Proceedings of the Sixth IEEE International Conference on Fuzzy Systems, 2 (1997) 1115-1120.
- Raj, P.A., and Kumar, D.N., Ranking alternatives with fuzzy weights using maximizing set and minimizing set, Fuzzy Sets and Systems, 105 (1999) 365-375.
- Richei, A., Hauptmanns, U., and Unger, H., The human error rate assessment and optimizing system HEROS-a new procedure for evaluating and optimizing the man-machine interface in PSA, Reliability Engineering and System Safety, 72 (2001) 153-164.
- Ruspini, E.H., Bonissone, P., and Pedrycz, W., (eds.), Handbook of Fuzzy Computation, Institute of Physics Publishing, Bristol and Philadelphia, 1998.
- Shen, X.X., Tan, K.C., and Xie, M., The implementation of quality function deployment based on linguistic data, Journal of Intelligent Manufacturing, 12 (2001) 65-75.
- Sii, H.S., and Ruxton, T., and Wang, J., A fuzzy-logic-based approach to qualitative safety modelling for marine systems, Reliability Engineering and System Safety, 73 (2001) 19-34.
- Steimann, F., and Adlassnig K.P., A fuzzy medical data model, Proceedings of the 12th European meeting on Cybernetics and Systems Research, 1 (1994) 271-278, World Scientific, Singapore.
- Tanaka H., and Guo, P., Possibilistic Data Analysis for Operations Research, Physica-Verlag, Heidelberg, New York, 1999.
- Tanaka, H., and Lee, H., Interval regression analysis by quadratic programming, IEEE Transactions on Fuzzy Systems, 6 (1998) 473-481.
- Tanaka, H., Uejima, S., and Asai, K., Fuzzy linear regression model, in International Congress on Applied Systems and Cybernetics, Acapulco, Mexico, VI, (1980) 2933-2938.
- Tanaka, H., Uejima, S., and Asai, K., Linear regression analysis with fuzzy model, IEEE Transactions Systems Man and Cybernetics, 12 (1982) 903-907.

- ten Berge, J.M.F., Least squares optimization in multivariate analysis, DSWO Press, Leiden, 1993.
- Watada, J., and Yabuuchi, Y., Fuzzy principal component analysis and its application, Biomedical Fuzzy and Human Sciences, 3 (1997) 83-92.
- Xu, R., and Li, C., Multidimensional least-squares fitting with a fuzzy model, Fuzzy Sets and Systems, 119 (2001) 215-223.
- Yabuuchi, Y., Watada, J., and Nakamori, Y., Fuzzy Principal Component Analysis for Fuzzy Data, in: Proceedings of the Sixth IEEE International Conference on Fuzzy Systems, 2 (1997) 1127-1132.